

Association for Information Systems AIS Electronic Library (AISeL)

Wirtschaftsinformatik Proceedings 2003

Wirtschaftsinformatik

September 2003

Analytische Informationssysteme im Energiehandel

Carsten Felden

Universität Duisburg-Essen, C.Felden@uni-duisburg.de

Follow this and additional works at: <http://aisel.aisnet.org/wi2003>

Recommended Citation

Felden, Carsten, "Analytische Informationssysteme im Energiehandel" (2003). *Wirtschaftsinformatik Proceedings 2003*. 76.
<http://aisel.aisnet.org/wi2003/76>

This material is brought to you by the Wirtschaftsinformatik at AIS Electronic Library (AISeL). It has been accepted for inclusion in Wirtschaftsinformatik Proceedings 2003 by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

In: Uhr, Wolfgang, Esswein, Werner & Schoop, Eric (Hg.) 2003. *Wirtschaftsinformatik 2003: Medien - Märkte - Mobilität*, 2 Bde. Heidelberg: Physica-Verlag

ISBN: 3-7908-0111-9 (Band 1)

ISBN: 3-7908-0116-X (Band 2)

© Physica-Verlag Heidelberg 2003

Analytische Informationssysteme im Energiehandel

Carsten Felden

Universität Duisburg-Essen

Zusammenfassung: Die Marktliberalisierung zwang Energiedienstleistungsunternehmen, neue Informationssysteme einzuführen, um Energiehändler bei analytischen Aufgaben zu unterstützen. Neben dem klassischen Ansatz, mittels Data Warehouse zeitbezogene Einblicke in das Marktgeschehen zu geben, ist es entscheidend, zusätzlich externe Informationen aus dem Internet verfügbar zu machen. Wetterinformationen, politische Nachrichten oder Marktgerüchte sind erforderlich, um die Variablen des volatilen Energiemarktes richtig zu interpretieren. Ausgehend vom multidimensionalen Datenmodell und erfassten Markttransaktionen wird eine Datenbank aufgebaut, die Energiehändler analytisch unterstützt. Zusätzlich gilt es, externe Informationsquellen zu finden und deren Informationen nach einem Filterungsprozess im Data Warehouse zu erfassen. Diese qualifizierten Informationen werden, über die Zeitachse mit Marktdaten verknüpft, in einer zentralen Benutzerschnittstelle dargestellt.

Schlüsselworte: Data Warehouse, Energiehandel, Externe Daten, Datenfilterung, Marktdateninformationssystem

1 Der Energiemarkt in der Bundesrepublik Deutschland

Die Informationsversorgung zur Entscheidungsunterstützung und -bereitstellung erweist sich in vielen Unternehmensbereichen heute noch als unbefriedigend [Stre96]. In jüngster Zeit konzentrierten sich die Forschungsanstrengungen zunehmend auf die Erweiterung der zentralen Datenhaltung durch unternehmensexterne Daten. Neue Kommunikationsmedien und insbesondere das Internet gehören zu den beherrschenden Themen der öffentlichen Diskussion. Das exponentielle Wachstum des Internet führt zu einem immer größer werdenden Angebot von Daten, die eine Entscheidungsunterstützung darstellen können. Für den Entscheidungsträger hat dies eine steigende Komplexität und eine abnehmende Transparenz zur Folge. Aus diesem Grunde ist eine Zusammenführung unternehmensinterner und -externer Daten zur Entscheidungsunterstützung von großer Bedeutung. Dies gilt insbesondere für den liberalisierten Energiemarkt. So ist es beispielsweise

se bereits heute schwierig, unter der Vielzahl von Stromanbietern, Strombörsen und Strombrokern denjenigen zu finden, der ein bestimmtes Produkt zu den günstigsten Konditionen anbietet. Selbst mit derzeit vorhandenen effizienten Werkzeugen wird es schwierig sein, die zukünftig zu erwartende Komplexität des Energiemarktes in einem Marktdateninformationssystem abzubilden, um im Bereich des Energiehandels aktiv agieren zu können.

Spezifische Datenbanksysteme, die eine besondere Problemadäquanz für diesen Aufgabenbereich versprechen, gewinnen daher in letzter Zeit zunehmend an Bedeutung. Gerade bei der Gestaltung solcher Informationssysteme zeigt sich, dass für die Datenbankmodellierung, die Datenintegration sowie die Datenpräsentation keine durchgängige Lösung verfügbar ist. Die am Markt bekannten Konzepte verfügen nur über Lösungen zu Teilbereichen der Informationssystementwicklung, ein integriertes Konzept ist noch nicht auszumachen. Konzepte zur Informationserfassung und -bereitstellung sind zwar vorhanden, jedoch für die Problemstellung nicht ausreichend. Der Informationsbezug aus Quellen, deren Struktur und Inhalte bekannt sind, ist inzwischen mit Standardlösungen beherrschbar.¹ Schwieriger hingegen gestaltet sich das Auffinden und Überprüfen von Daten aus unbekannten Quellen in der großen unübersichtlichen Verfügbarkeit von Informationen. Zudem liegen die Daten in der Regel meist in einer Formatierung vor, die eine automatische Weiterverarbeitung im Unternehmen nicht möglich macht. Somit besteht die Anforderung, ein Konzept zu gestalten, das eine Identifikation relevanter, externer Daten aus ggf. unbekannten Quellen ermöglicht, um diese den Kompetenzträgern verfügbar zu machen [Raa⁺00, S. 55 f.].

Diese heterogenen Daten² erfordern eine entsprechende Benutzeroberfläche. Es wird zusätzlich dargestellt, inwieweit das Konzept der Visualisierung helfen kann, den Energiehändlern die Informationsvielfalt sinnvoll und nachvollziehbar aufzubereiten, um tatsächliche Entscheidungsunterstützung, wie diese auch vom Data-Warehouse-Konzept gefordert wird, zu gewährleisten. Die Entwicklung einer solchen Benutzeroberfläche wird im Rahmen dieses Beitrages skizziert.

2 Einsatz von Informationssystemen im Energiehandel

Obwohl Marktdaten für den Preisbildungsprozess von essentieller Bedeutung sind, besitzen nur die wenigsten Unternehmen ein entsprechend implementiertes Infor-

¹ Die MS-Office-Produkte bieten beispielsweise inzwischen Lösungen, unterschiedliche Datenformate zu importieren.

² Im Internet befindet sich eine Vielzahl unterschiedlicher Datenformate, die in dieser Ausprägung in Unternehmen nicht zu finden ist.

mationssystem für Energieanalysten. Viele Unternehmen sind noch damit beschäftigt, entsprechende Kommunikationsstrukturen innerhalb ihres Unternehmens zu schaffen, so dass ein solches System eher nachrangige Bedeutung aufweist. Dies gilt sowohl für große Energiedienstleistungsunternehmen als auch für kleine Stadtwerke. Aus diesem Grunde weichen viele Analysten auf das Programm Excel der Firma Microsoft aus. Dieses sicherlich sehr leistungsstarke Programm hat allerdings den Nachteil, dass Informationen nicht kontinuierlich über einen langen Zeitraum gesammelt werden, die allen interessierten Mitarbeitern gegebenenfalls unternehmensweit zur Verfügung gestellt werden können. So besteht die inhärente Gefahr, dass identische Analysen zu unterschiedlichen Zeiten nicht vergleichbar sind und darüber hinaus, dass Wissen, z. B. bei einem Mitarbeiterwechsel, in einem Unternehmen verloren gehen kann.

Jedoch wurden in der Zwischenzeit zwei Systeme entwickelt, die zur Unterstützung der Energieanalysten Marktdaten sammeln und entsprechend zur Verfügung stellen. Zum einen ist dies das Systematic Analysis and Research Tool (SMART), dass im Jahr 2000 gemeinschaftlich von der damaligen VEW Energie AG und einer universitären Einrichtung entwickelt wurde. Dies ist von der RWE Energie AG übernommen worden und den gestiegenen Bedürfnissen des wachsenden Marktes angepasst worden. Auf Grund der Projekterfahrung kann dieses System an dieser Stelle aus Projektsicht beschrieben werden. Zum anderen ist es das Marktinformationssystem für Analyse und Planung (MESAP) der Firma sevenZone, die als einziger Hersteller derzeit eine Lösung auf dem Markt anbieten. Dieses Tool ist u. a. bei der EnBW AG im Energiehandel im Einsatz und wird im Rahmen dieses Artikels dem Tool SMART gegenübergestellt.

2.1 SMART - Systematic Analysis and Research Tool

Anhand der Kenntnisse [Stah95, S. 259] über die Tätigkeiten der Anwender wurden allgemein formulierte Szenarien erstellt, welche die Basis für den Aufbau der Berichte des Marktdateninformationssystems bilden. Die einzelnen Szenarien wurden zu Themengruppen, zusammengefasst. Diese Gruppen bilden auch die späteren multidimensionalen Datenstrukturen, von denen insgesamt sechs abgebildet worden sind. Durch die Erstellung repräsentativer Aufgaben kann zunächst ein gegenseitiges Verständnis und bereits ein frühzeitiger Systemprototyp geschaffen werden [Niel93, S. 100]. Im SMART-Projekt wurden 29 verschiedene repräsentative Aufgaben identifiziert, die anhand von sechs Berichten realisierbar sind. Einer dieser Berichte ist im Folgenden exemplarisch dargestellt. Parallel zum bisherigen Vorgehen wurden ein Glossar und ein Data Dictionary [Wert86] aufgebaut, um sich zum einen auf die Begriffswelt des Energiehandels einzustimmen und zum anderen die Strukturen des Analytischen Informationssystems zu dokumentieren.

Für die Abbildung der Ergebnisse der Informationsbedarfsanalyse wurde das Application Design for Analytical Technologies (ADAPT) [Bulo98, S. 253 ff.] als Modellierungsmethode gewählt. Diese Entscheidung resultiert aus den Unzulänglichkeiten des Entity Relationship Modelling (ERM) [Chen76] bei der Darstellung multidimensionaler Sachverhalte. Da eine relationale Datenbank [Elm⁺94] zur Verfügung stand, wurde die multidimensionale Datenstruktur in Form eines Star Schemas [Hahn99, S. 152 ff.] realisiert. Im Rahmen des Ermessensspielraums, den ADAPT bei der Verwendung der Beschreibungselemente bietet, wurde darauf geachtet, möglichst wenige unterschiedliche Symbole zu verwenden. Da das Datenmodell im Rahmen des Prototyping [Balz01, S. 100 ff.] intensiv zum Gedankenaustausch mit dem fachlichen Projektleiter genutzt werden sollte, war eine verständliche Darstellungsweise geboten. Die folgende Abbildung 1 zeigt den multidimensionalen Datenraum *Product Price History* in der ADAPT-Notation.

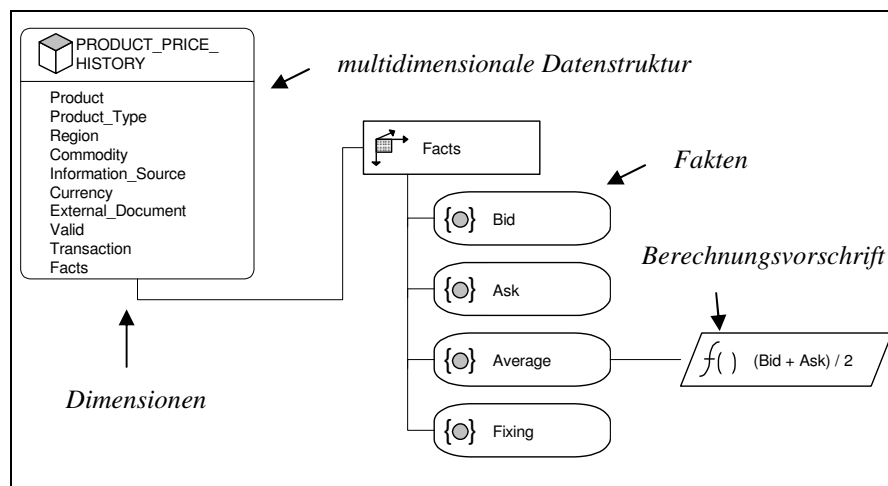


Abbildung 1: Multidimensionaler Datenraum zum Bericht *Product Price History*

Ziel dieser Analyse *Product Price History* ist, die jeweiligen Bid-, Ask- (inklusive deren Durchschnitt)³ und Fixingkurse⁴ der Stromhändler im Zeitverlauf darzustellen. Dabei können sich die Kurse auf verschiedene Regionen (z. B. CEPI, EIS)⁵ oder auf verschiedene Produktarten (z. B. Base oder Peak) beziehen. Der Bericht wird durch das Beschreibungselement *Hyperwürfel* gekennzeichnet. Im oberen Bereich des Symbols steht die Bezeichnung, im unteren Bereich werden sämtliche Dimensionen aufgelistet, welche den multidimensionalen Datenraum

³ Bei Bid und Ask handelt es sich um Kauf- und Verkaufgebote.

⁴ Das Fixing ist der jeweilige Schlusskurs eines Handelstages.

⁵ CEPI bezeichnet Norddeutschland; EIS bezeichnet Süddeutschland.

aufspannen und die darin enthaltenen Fakten beschreiben. Daraus resultiert folgendes:

- Die Dimensionsbezeichnungen verweisen auf die gleichnamigen Dimensionsdarstellungen dieses Modells;
- *External_Document* verweist auf die Dimension, in der die Internetdokumente erfasst sind;
- *Valid* und *Transaction* beziehen sich beide auf die Dimension *Time* [Stoc01]. *Valid* ist hierbei die Gültigkeitszeit und *Transaction* die Transaktionszeit. Diese sind eingeführt worden, damit bei Fehlerkorrekturen der jeweils gültige Wert anhand der aktuellsten Transaktionszeit festgestellt werden kann. Aus Gründen der Revisionsicherheit bleiben falsch eingegebene Werte in der Datenbank gespeichert;
- Lediglich die Kennzahlen (Facts) werden explizit abgebildet. Unter Verwendung des Symbols *Dimensionselement* werden die für diesen Bericht relevanten Fakten aufgezählt;
- Das Faktum *Average* wird aus den Fakten *Bid* und *Ask* berechnet. Daher ist es mit dem Beschreibungselement *Berechnungsformel* verbunden, in dem die Berechnungsvorschrift angegeben ist.

Um für sämtliche Berichte den multidimensionalen Datenraum aufspannen zu können, wurden insgesamt 14 verschiedenen Dimensionen identifiziert. Diese sind *Product*, *Product_Type*, *Option_Style*, *Option_Type*, *Region*, *Commodity*, *Price_Index*, *Information_Source*, *External_Document*, *Currency*, *Time*, *Interest_Rate*, *Energy_Production* und *User*. Aus den Datenwürfeln gingen die Faktentabellen hervor. Da mehrere Faktentabellen vorliegen, handelt es sich beim logischen Datenmodell der SMART-Datenbank um eine Galaxie [Hahn99, S. 158 f.] mit 1:n-Beziehungen zwischen den Dimensions- und Faktentabellen.

Die Implementierung des oben beschriebenen Datenmodells erfasst Daten aus internen sowie bekannten externen Quellen. Externe Quellen sind zunächst Datenströme externer Informationsanbieter, mit denen ein entsprechender Vertrag geschlossen wurde. Hierbei handelt es sich z. B. um den *Deutschen Wetterdienst* sowie On-Line-Broker wie *Prebon* oder *Intergrid*. Durch diese Datenströme lassen sich Zeitreihen zu Betriebs- und Marktdaten, z. B. Schlusskurse börsengehandelter Produkte, darstellen. Ebenso werden Kraftwerksdaten als Stamm-, Auslastungsdaten sowie Kraftwerksausfälle gespeichert. Wetterdaten ermöglichen den Rückschluss auf Schwankungen im Stromverbrauch der Kunden in bestimmten Regionen.

2.2 MESAP – Modulare Energiesystemanalyse und Planung

Das System MESAP unterscheidet sich in der Art der erfassten Daten nicht vom SMART-System. Grundlegendes Ziel von MESAP ist ebenso die Sammlung relevanter Daten, Rohdaten zu validieren und zu dokumentieren, Daten zu analysieren und relevante Aussagen aufzubereiten sowie die Informationen zu kommunizieren und zu publizieren. Abbildung 2 zeigt die Architektur von MESAP [seve03].

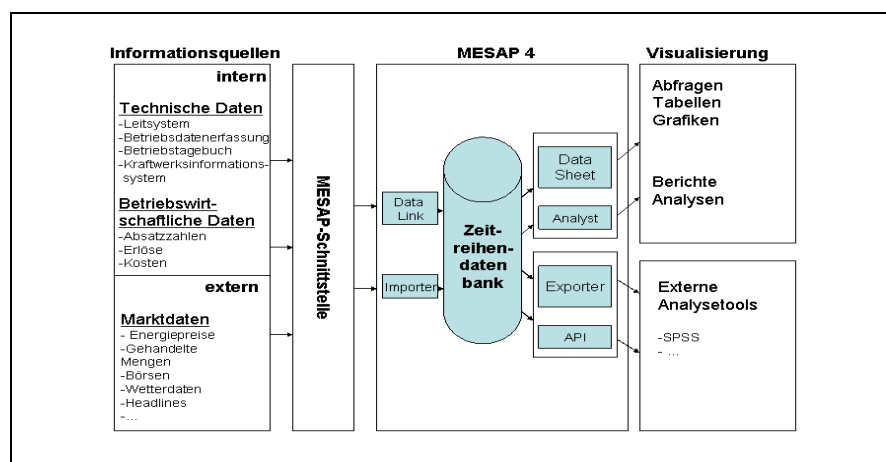


Abbildung 2: MESAP-Architektur [seve03]

MESAP besteht aus vier Bauelementen. Das Herzstück bildet die Datenbank, die mit multidimensionalen Strukturen in einer relationalen Datenbank implementiert ist. Innerhalb der Dimensionen, die den multidimensionalen Datenraum aufspannen, besteht die Möglichkeit zur Aggregation der erfassten Zeitreihendaten. Das zu Grunde liegende Datenmodell ist jedoch kein Star Schema, wie es in der einschlägigen Literatur beschrieben wird. Vielmehr orientiert sich der Aufbau der Datenbank an der intuitiven Umsetzung der geforderten Berichte der Anwender durch die Entwickler [seve03].

Die Daten lassen sich aus unterschiedlichen Quellen importieren, was mittels des zweiten Bausteins, der MESAP-Schnittstelle, geschieht. Dabei können zurzeit externe Daten lediglich über ein Standardformat eingelesen werden. Zugriff auf diese Daten ist mittels der Benutzerschnittstelle, dem dritten Baustein, möglich. Dort lassen sich Datensichten konfigurieren und speichern sowie ad-hoc-Analysen durchführen. Grundsätzlich wird Microsoft Excel als Bericht- und Analysetool verwendet, da es, wie bereits angesprochen, viele Analysefunktionen enthält und den Anwendern aus der täglichen Arbeit vertraut ist. Dabei ist aber auch der Einsatz weiterer Analysetools wie z. B. SPSS über eine entsprechende Schnittstelle möglich [seve03].

3 Weiterentwicklung des SMART-Systems

In der universitären Forschung konzentrierte sich die Weiterentwicklung darauf, externe Daten aus bisher unerschlossenen Quellen zu finden, zu extrahieren und in der Datenbank zu erfassen. Dies ist in noch keinem bestehenden System realisiert. Zusammenhängend damit wurde die Schwierigkeit erkannt, heterogene Daten in einer Benutzeroberfläche zu präsentieren. Aus diesem Grunde wurde das Konzept der Informationsvisualisierung als Lösungshilfe untersucht.

3.1 Konzepte zur Informationserfassung und -bereitstellung

Es ist sicherlich kein neuer Ansatz, Informationen für Führungs- und Entscheidungsträger zur Verfügung zu stellen, da schon frühzeitig erkannt wurde, dass operative Informationssysteme keine wirkliche Unterstützung bei Analyseaufgaben geben können. Unter dem Begriff *Management Support Systeme* werden schon seit Jahren Softwaresysteme subsumiert, welche diese Aufgabe leistungsfähig erfüllen können [Glu⁺97, S. 39 ff.]. Insbesondere das Data-Warehouse-Konzept hat sich mit seinen organisatorischen Prinzipien und Zielen zur Erfassung und Bereitstellung von Massendaten durchgesetzt [Inm⁺00]. Damit bildet dieses Konzept auch die Grundlage des zu entwickelnden Marktdateninformationssystems. Jedoch konzentriert sich das Data-Warehouse-Konzept primär auf unternehmensinterne Daten. Aus diesem Grunde wurde versucht, das Data Warehouse durch die Integration externer Daten zu erweitern. Der erste Ansatz hierzu stammt von Richard D. Hackathorn, das sogenannte *Web Farming for the Data Warehouse* [Hack98]. Das Erfordernis, interne Daten mit externen Daten anzureichern, wurde nicht nur von Hackathorn erkannt.⁶ Unter anderem wird auch der Datenbestand des Redaktionsleitstandes der SAP AG um externe Daten erweitert, um eine vollständige Datenbasis für Entscheidungen anbieten zu können.

Grundsätzlich versucht das Web Farming durch die vier organisatorischen Schritte Getting Started, Getting Serious, Getting Smart und Getting Tough, die erfassten Daten eines Data Warehouse um diejenigen aus dem Internet zu erweitern [Hack98, S. 10 ff.]. Dadurch wird den wandelnden Bedingungen einer dynamischen Umwelt Rechnung getragen. Dabei geht es nicht um orientierungsloses Internetsurfen mit Sondierung der Suchergebnisse, sondern um einen evolutorischen und systematischen Prozess zur Filterung von Datenquellen im Internet, um geschäftskritisches Wissen zu erfassen.

Der Redaktionsleitstand ist ein weiteres Konzept, Managern entscheidungsrelevante Daten zur Verfügung zu stellen. Er ist nicht als ein Steuerungspult für Manager zu verstehen [Meie00, S. 3 ff.]. Wie in jedem anderen Projekt zur Erstellung analytischer Datenbanken gilt es zunächst, den Datenbedarf der Entscheidungsträ-

⁶ Weiter Ansätze sind auch bei [Teut01] [Uhr⁺98] nachzulesen.

ger zu identifizieren, diesen dazu zu nutzen, um gezielt führungsrelevante Daten im Internet zu suchen. Nutzbare Internetdaten und interne Daten werden miteinander verknüpft, um dem Entscheidungsträger interne quantitative und externe qualitative Daten zur Verfügung zu stellen. Der Redaktionsleitstand soll sämtliche Informationsversorgungsphasen von Führungskräften unterstützen. Aus diesem Grunde kommt ein Prozess zum Einsatz, der in Form einer Wertschöpfungskette mit der Informationsbedarfsermittlung beginnt und mit einer Datenpräsentation abschließt.

3.2 Integrationsprozess externer Daten aus unbekannten Quellen

In weiteren universitären Forschungsaktivitäten wird auf Grund der vielen manuellen Eingriffe bei den zuvor aufgezeigten Konzepten angestrebt, die Datenbasis automatisiert mit externen Daten zu erweitern. Ziel ist, diese Quellen im World Wide Web (WWW) zu finden und den Analysten zur Verfügung zu stellen. Die folgende Abbildung 3 zeigt den Verlauf, so wie er in einem universitären Forschungsprojekt vorgeschlagen wird [Feld02, S. 57 ff.].

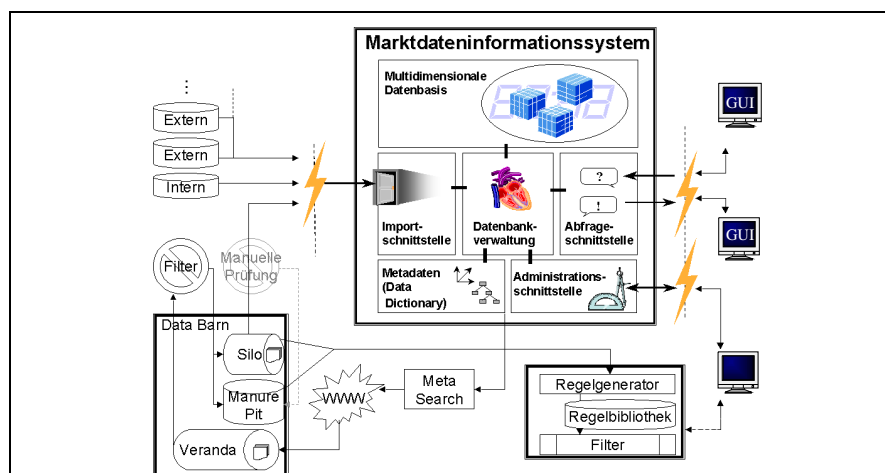


Abbildung 3: Architektur des Marktdateninformationssystems SMART

Aufbauend auf der bereits vorgestellten SMART-Architektur werden Metadaten aus dem Data Dictionary ausgewählt und als Suchbegriffe für eine Suche im WWW (Meta Search) eingesetzt. Die Suchergebnisse werden in die Data Barn übertragen. Die Data Barn stellt transiente (Veranda und Silo) und persistente (Manure Pit) Speicher für die Bearbeitung der Internetseiten zur Verfügung. Den zentralen Problembereich der Integration bildet in diesem Kontext die Regelgenerie-

rung. Zur Klassifizierung wendet ein Filter die zuvor generierten Regeln an. Dazu wird ein Multilayer Perceptron (MLP) [Bish95, S. 117 ff.] [Khan90, S. 50] mit Deskriptoren, Deskriptorenergänzungen sowie Datenqualitätsanforderungen angewendet.⁷ Zum Abschluss des Integrationsprozesses müssen die interessanten Internetseiten mit den jeweiligen multidimensionalen Datenstrukturen verbunden werden, um bei den Analysen den Zugriff zu ermöglichen. Dies geschieht in Form einer weiteren Dimension für diese Datentypen im Star Schema.⁸ Die uninteressanten Internetseiten verbleiben als Trainingsdatenbestand in der persistenten Manure Pit.

3.3 Integration externer Daten

Für die Klassifizierung von Internetdokumenten ist insbesondere die Informationsidentifikation und Informationsbewertung, die durchgeführt werden muss, von Interesse. Da für die Informationsidentifikation eine Metasuchmaschine eingesetzt wird, konzentriert sich der folgende Abschnitt auf die Informationsbewertung. Grundsätzlich konzentriert sich die Bewertung auf Textdokumente, da eine automatische inhaltliche Bewertung von Bild-, Audio- und Videodateien zurzeit nicht ohne Schwierigkeiten durchführbar ist. Zur Bewertung ist der Aufbau des Regelgenerators ein zentrales Problem. Es sind Regeln zu generieren, die in einem Filter ausschließlich diejenigen Internetdaten für das Marktdateninformationssystem identifizieren, die auch tatsächlich für Entscheidungsträger relevant sind.

3.3.1 Deskriptorenbestimmung

Um Attribute für eine Klassifikationsstruktur auswählen zu können, müssen sowohl Anforderungen an die Daten operationalisiert als auch Methoden des Text Mining berücksichtigt werden, um die interessanten Seiten für Kompetenzträger zu bestimmen und zur Verfügung stellen zu können. Die Attribute der Klassifikationsstruktur bilden das sogenannte *Informationssurrogat*. Dabei besteht dieses Surrogat aus einem Vektor, der eine Teilmenge der Metadaten, Qualitätsanforderungen sowie begriffliche Erweiterungen erfasst. Für das Marktdateninformationssystem wird davon ausgegangen, dass Textdaten aus dem Internet im Unternehmen verfügbar sind. Diese müssen entsprechend manuell als *interessant* und *uninteressant* klassifiziert werden. Mit den vorhandenen Seiten wird eine Trainingsbasis gebildet, um die Deskriptoren zu identifizieren, die interessante und uninteressante Texte charakterisieren.

⁷ Zu beachten ist, dass zurzeit das Ergebnis der Filterung manuell zu überprüfen ist.

⁸ Aus diesem Grunde ist eine objektrelationale Datenbank [Küng94] erforderlich.

3.3.2 Informationsqualität

Ein wichtiger Erfolgsfaktor für die Nutzung von Informationssystemen ist die Qualität der gespeicherten Daten. Betriebliche Entscheidungen auf allen Ebenen basieren auf Informationen, die aus diesen Daten gewonnen sein können. Somit ermöglichen es nur Daten guter Qualität Information zu erzeugen, die als Basis für zuverlässige und sachgerechte Entscheidungen dienen können [Naye93, S. 51 ff.]. Im Zusammenhang mit der Nutzung eigenständiger Informationssysteme für Analysezwecke tritt dieser Aspekt in den Vordergrund. Ein zunehmender Automatisierungsgrad in den Analyseprozessen führt dazu, dass die durch das Erfahrungswissen des Entscheiders bedingte gedankliche Korrektur falscher Daten so nicht mehr erfolgen kann. Durch das Internet entstand nicht nur eine quantitative Steigerung des Informationsangebotes, analog dazu wuchsen auch die Unterschiede in der inhaltlichen und grafischen Qualität. Um die Quellen einordnen zu können, wird der jeweilige Verfasser betrachtet. Dessen Organisationsform und Finanzierungsform (z. B. Werbeeinnahmen) ist hierzu entsprechend zu analysieren. Die Qualität der Quelle wird anhand der drei Kriterien Korrektheit, Aktualität und Navigation bewertet. Aufbauend auf diesen Kriterien kann eine Rangfolge der Internetquellen erstellt werden, wobei subjektive Einflüsse idealerweise nicht einfließen. Die Reihenfolge entsteht durch die quantitative Gewichtung ausgewählter Kriterien [Raa⁺00, S. 56 f.]. Zentraler Problembereich der aufgeführten Qualitätsanforderungen ist, diese für Informationen zu operationalisieren, damit sie auch in einem Filter automatisch genutzt werden können. Für das Marktdateninformationssystem bedeutet das, dass geprüft wird, ob ein Kriterium genannt ist oder nicht. Hierzu werden die Eigenschaften eines Dokumentes genutzt, in denen die Informationen erfasst sind. Die Gewichtung der Qualitätsfaktoren ist Bestandteil des Regelgenerierungsprozesses.

3.3.3 Fachspezifische Deskriptorenergänzungen

Für den Prozess ist es sinnvoll, die identifizierten Deskriptoren eines Textes um Begriffe zu erweitern, die eine Klassenzuordnung verbessern. Hintergrund hier ist, dass Seiten anhand der bisherigen Kriterien als interessant eingestuft werden können, die dem Kompetenzträger keinen tatsächlichen Informationsmehrwert generieren. Dies bedeutet, dass weitere Charakteristika, die den Unterschied in der Interessantheit ausmachen, aber nicht in den ermittelten Deskriptoren enthalten sind, hinzugefügt werden müssen. Dies wird manuell geschehen müssen, da dieser Unterschied vom Benutzerkreis des Systems festgelegt und nicht automatisch erkannt wird. Wäre dies möglich, müsste keine Deskriptorenergänzung stattfinden, da diese bereits vom Regelgenerator erkannt würde.

3.4 Informationsfilter

Die Filterung beschreibt die Anwendung der zuvor generierten Regeln, die in einer Regelbibliothek erfasst sind. Interessante Daten sind auszuwählen und dem Benutzer zur Verfügung zu stellen. Durch ein Künstliches Neuronales Netz [McC⁺43] [vonN45] [Rand82] [McC⁺90] [Zaun99] [Bene98] [Brau95] [Kinn94] [Mazz92] [Kinn94] kann eine Klassifikation durchgeführt werden. In diesem Fall handelt es sich bei dem Neuronalen Netz um ein Multilayer Perceptron (MLP). Die folgenden Abbildung 4 zeigt das MLP.

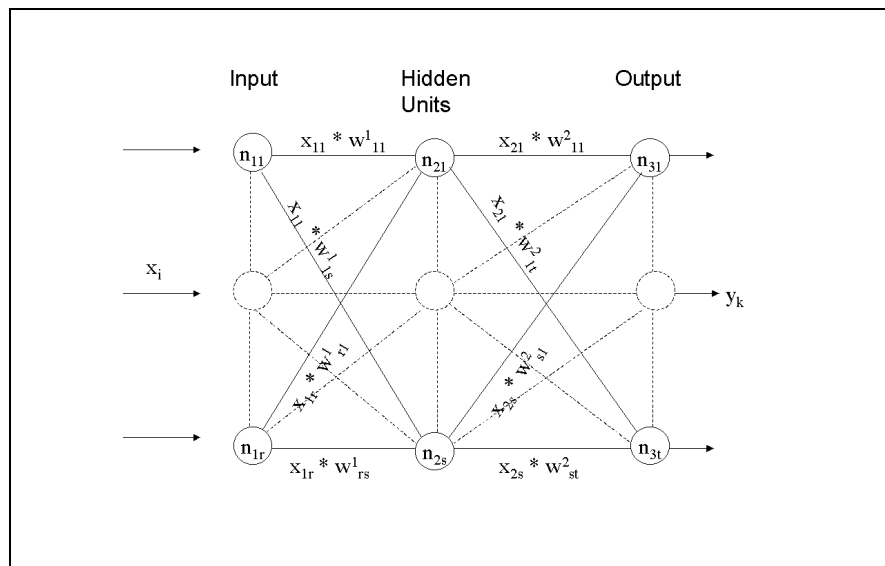


Abbildung 4: Multilayer Perceptron zur Textklassifikation

Das dargestellte Netzwerk führt zu einer Transformation der Werte durch zwei nachgelagerte Ebenen. Die Abbildung enthält r Inputneuronen (n_{11}, \dots, n_{1r}), s Hidden Units (n_{21}, \dots, n_{2s}) und t Outputneuronen (n_{31}, \dots, n_{3t}), mit denen der Neuronenwert x_i verbunden ist [Bish95, S. 116 ff.].

Die Inputschicht wird durch den Abgleich der einzelnen Deskriptoren, Qualitätsanforderungen und Ergänzungen gebildet. Diese bilden die Eingabe zur versteckten Schicht. Hier werden, getrennt nach Art des Neuron, die ankommenden Signale mit der Eingangsstärke w_{ij} multipliziert. Die Netzeingabe net_{2j} wird aus der Summe der gewichteten Inputs bestimmt. Aus net_{2j} wird mit der Aktivierungsfunktion $a_{2j}(net_{2j})$ die Aktivierung berechnet. Die drei Einheiten der versteckten Schicht geben ihren Output als Eingabe an die Ausgabeschicht weiter. Dort erfolgt erneut eine Aktivierung durch die Schrittfunktion sowie eine Umrechnung in

die Ausgabe, die letztendlich die Klassifikation darstellt. Der Multiplikator der Ausgabefunktion liegt im Einheitsintervall in $\Re [0, 1]$.

Nach Anwendung des Neuronalen Netzes ist das Ergebnis unter Beachtung eines festgelegten Schwellenwertes eine Klassifikation der relevanten Internetseiten in *interessante* (positiv, wenn oberhalb des Schwellenwertes) und *uninteressante* (negativ, wenn unterhalb des Schwellenwertes) Seiten. Die Datensuche und anschließende Datenbewertung durch den Filter lässt sich zusammenfassend auch als ein stationärer Informationsagent, so wie ihn Zarnekow beschreibt, bezeichnen [Zarn99, S. 226 f.].⁹

3.5 Nutzung des Visualisierungskonzeptes für eine einheitliche Benutzeroberfläche

Ein zentrales Problem besteht darin, dem Analysten die Daten aus heterogenen Quellen in einer Oberfläche zur Verfügung zu stellen, so dass dieser eine tatsächliche Entscheidungsunterstützung erfährt.

Dass die Visualisierung Dokumente und deren Beziehung zueinander darstellen soll, erscheint zunächst nicht überraschend. Diese Forderung besagt jedoch, dass dem Anwender die Problemdomäne abstrakt bewusst gemacht wird. Die Visualisierung ist vergleichbar mit einem Tor, durch das ein Anwender in den Datenraum eintritt, um mit den vorhandenen Daten oder Dokumenten Analysen durchzuführen. Dabei sind die darzustellenden Beziehungen die Relationen zu den Dimensionen, welche den Datenraum der Dokumente aufspannen. So wird der Kontext aufgezeigt, in dem sich die einzelnen Dokumente befinden. Es besteht beispielsweise eine Beziehung zur Zeitdimension und zum Rating der einzelnen Dokumente. Aus dem Schnittpunkt beider resultiert die Positionierung des Dokumentes im Diagramm. Die gewählte Darstellungsform soll dem Anwender den Informationsgehalt einzelner Dokumente als Ergänzung zur jeweiligen multidimensionalen Datenstruktur verständlich machen. Als gut bewertete Internetdaten werden in der Darstellung physikalisch höher angeordnet als schlecht bewertete Daten.

Die Navigation ist eine weitere Anforderung an die Visualisierung. Navigation bietet dem Anwender die Möglichkeit, in der Visualisierung auf eine Art Entdeckungsreise zu gehen. Da die Darstellung gleichzeitig das Rating der Informationen widerspiegelt, soll der Anwender die Möglichkeit haben, das Rating oder den betrachteten Zeitausschnitt jederzeit verändern zu können. Zudem sollen Änderungen in der Auswahl der Dimensionsattribute für die Darstellung der Marktinformationen gleichzeitig die Auswahl der angezeigten Internetdokumente beeinflussen. Da beim *onMouseover-Effekt* der Quell-Uniform-Ressource-Locator

⁹ Grundlegende Arbeiten finden sich bei [Woo⁺94]. Eine abweichende Darstellung Intelligenter Softwareagenten wählt [Gehr00].

(Quell-URL) von Internetdaten angegeben werden kann, soll diese URL auch als Hyperlink verfügbar sein, um gegebenenfalls Aktualisierungen direkt überprüfen zu können. Zusätzlich erscheint es sinnvoll, Sprünge zwischen inhaltlich verwandten Dokumenten direkt zu ermöglichen.

Die Steuerung der Navigation soll vom Anwender lediglich durch Nutzung der Maus erfolgen. Das Erlernen und Anwenden von Befehlswörtern zur Navigation ist unbedingt zu vermeiden. Durch den bereits angesprochenen *onMouseOver*-Effekt besteht die Möglichkeit, dem Anwender kurze und prägnante Informationen zu den Navigationspunkten zu liefern. Würden jeweils Befehlswörter eingegeben werden, verlängerte sich der Zeitraum, in dem ein Anwender ein System erlernt¹⁰, und ein dynamisches Arbeiten und Navigieren würde behindert. Um die Systemsteuerung durch den Anwender zu erleichtern, sollen Navigationshilfen permanent visuell auf der Oberfläche abgebildet sein. Zu diesem Zweck bieten sich beispielsweise sogenannte Slide-Buttons an. Ähnlich einem Scrollbalken sind diese in die Oberfläche integriert, um z. B. den Betrachtungsausschnitt im gewählten Zeitintervall durch Hin- und Herschieben verändern zu können. Analog ist es möglich, durch zwei Slide-Buttons die Intervallgrenzen für die betrachtete Zeitperiode oder für das Rating der Internetdokumente zu bestimmen.

Die im Marktdateninformationssystem benötigte Visualisierung unterscheidet sich von den herkömmlichen Konzepten der Abfragegenerierung und dem Information Retrieval durch die

- Fähigkeit des schnellen Filterns (Reduktion der Ergebnismenge),
- progressive Veränderung der Suchparameter,
- durchgehende Veränderung der Ziele und
- visuelle Analyse der identifizierten Ergebnisse [Ahl*99, S. 244 ff.].

Dies führt zu den folgenden Punkten, die im Marktdateninformationssystem umgesetzt werden sollen:

- Aufteilung der Oberfläche in unterschiedliche Bereiche;
- Dynamic Query Filters: Abfrageparameter werden unverzüglich mit Einstellungen der Slider, Buttons etc. neu justiert;
- Starfield Display: Die Ergebnismengen sind kontinuierlich verfügbar und unterstützen die Betrachtung hunderter oder tausender Datenbankelemente;
- Tight Coupling: Komponenten der Abfrage sind dahingehend verbunden, dass sie keine Darstellungsvarianten bieten, aber progressive Neudefinitionen er-

¹⁰ Die Folge bei Nutzung von Befehlsworteingaben ist eine i. d. R. verlängerte Einarbeitungsphase in das System.

möglichen. Insbesondere das Ergebnis einer Abfrage kann jederzeit als Input einer neuen Abfrage dienen.

Abbildung 5 zeigt die Benutzeroberfläche, die sich dem Anwender nach Programmstart präsentiert.

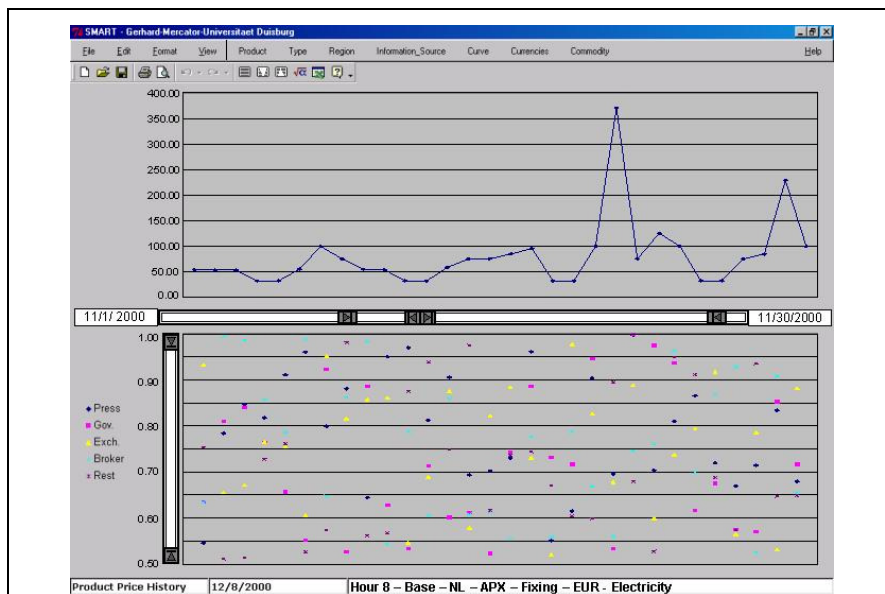


Abbildung 5: Benutzeroberfläche von SMART

Der geteilte Bildschirm wird optisch durch den mittig positionierten Timeslider erreicht. Das obere Diagramm stellt die Marktdaten aus bekannten Quellen in einem Kurvendiagramm dar, das untere Diagramm die in der Datenbank erfassten Internetdokumente in Form eines sogenannten Starfield Display [Car⁺99]. In diesem Kontext ist zu beachten, dass ein kausaler Zusammenhang zwischen den Kursverläufen und den ergänzenden Internetdaten nicht zwingend besteht. Beide Diagramme orientieren sich jeweils an der in der Datenbank erfassten Gültigkeitszeit (der Timeslider bezieht sich auf beide Darstellungen), dabei resultiert diese für die Internetdokumente aus dem Datum, an dem das Dokument auf dem jeweiligen Internetserver platziert wurde. In der Statusleiste rechts unten ist zu erkennen, welche Selektion in den Dimensionsmenüs vorgenommen wurde. Die möglichen Vollbilddarstellungen werden für Analysen der Energiehändler verwendet. Die rechte und linke Seite des Timeslider zeigt die Einstellung zu Gunsten der unteren und oberen Grenze des abgebildeten Zeitintervalls. Die durch Slidebuttons veränderbaren Einstellungen werden angezeigt. Durch Verschiebung der Pfeilbutton wird das Datum herauf- oder entsprechend heruntergesetzt. Analog dazu werden auch Slidebuttons für die Bestimmung der Grenzen für die Internetdokumente

verwendet. Diese befinden sich direkt neben dem Diagramm der Dokumentenanzeige.

Während die Darstellungsform der Internetdaten nicht variiert, kann die der Berichte unterschiedlich sein. Gestaltungsgrundlage für diese sind, wie bereits dargestellt, die ermittelten Anforderungen der Analysten, die bestimmen, welche Repräsentationsform für sie geeignet ist. Zur Darstellung der Werte des Berichts *Product_Price_History* wurde von Seiten der Anwender ein klassisches Liniendiagramm gewünscht. In diesem Diagramm wird der Preisverlauf der jeweiligen Produkte angezeigt.

Forschungsarbeiten zu *Dynamischen Abfragen* zeigen deutliche Verbesserungen in der Abfragegenerierung und einen hohen Grad an Zufriedenheit, wenn ein Anwender direkt Abfrageparameter via Maus manipulieren kann [Ahl⁺99, S. 245]. Zusätzlich ist jede der Abfragekomponenten (Slider, Button etc.) gleichzeitig ein Filter, der die Anzahl der Elemente in der visualisierten Ergebnismenge reduziert. Die Abfrageparameter werden dazu mit *Und*-Verknüpfungen zur Dimensionskombination verbunden. Dies ist für den Anwender einfacher, als eine Sequenz von Einzelabfragen zu generieren.¹¹

4 Konzeptpositionierung

Das entwickelte Marktdateninformationssystem unterscheidet sich von der üblichen Informationssystementwicklung in den einzelnen Phasen von der Datensuche bis zur Datenpräsentation. Dabei wird für dieses Informationssystem auf verfügbare Technologien zurückgegriffen. Die Datensuche wird sowohl beim Web Farming als auch beim Redaktionsleitstand manuell durchgeführt. Beide Konzepte nutzen Suchmaschinen, die entweder mit Parametern aus einer entsprechend erfassten Liste oder durch Stichworteingabe eines Anwenders gestartet werden. Der Redaktionsleitstand nutzt zusätzlich die Möglichkeit, sich Daten von sogenannten Informationsdiensten automatisch zusenden zu lassen. Anders das Marktdateninformationssystem: losgelöst von Benutzeraktionen wird ein Suchstring zu einem bestimmten Zeitpunkt automatisch mit Parametern aus dem Data Dictionary an eine Suchmaschine übergeben, die dann die Suche entsprechend ausführt.

Sind entsprechende Daten gefunden worden, so werden diese beim Redaktionsleitstand und beim Web Farming i. d. R. durch manuelles Copy & Paste in die Datenbank übertragen. Für das Web Farming wird darauf hingewiesen, dass für gleichbleibende Quellen ein Automatismus erstellt werden kann, der die Daten-

¹¹ Auf Grund der Anwendung von Regeln der Mengenlehre in Datenbanken wird so nur eine Ergebnismenge erzeugt. Bei einer Sequenz von Einzelabfragen müssten die Ergebnismengen aufwendig miteinander verknüpft werden.

übertragung ausführt. Dieser Automatismus wird im Rahmen dieses Konzeptes als Softwareagent bezeichnet. Auch hier unterscheidet sich das Marktinformationssystem. Es erfolgt zunächst keine Bewertung der Internetdaten. Sämtliche Suchergebnisse der Suchmaschine werden in die Data Barn übertragen. Dazu wird der gleiche HTTP-Aufruf verwendet, den auch Internet-Browser durchführen, um Internetdokumente aus dem Netz zu laden. So wird der vollständige Quellcode in die eigene Datenbank geladen und kann entsprechend weiterverarbeitet werden. Dazu ist keine Benutzerinteraktion erforderlich, da dieser HTTP-Aufruf ebenfalls automatisch gestartet und ausgeführt werden kann.

Beim Web Farming erfolgt die Bewertung der Internetdaten manuell. Der Redaktionsleitstand automatisiert dies teilweise. Es erfolgt ein Abgleich mit einer Deskriptorenliste, wodurch das Ähnlichkeitsmaß bestimmt wird. Durch das Ähnlichkeitsmaß kann erkannt werden, zu welchen internen Berichten die Internetdokumente zugeordnet werden können. Dies findet i. d. R. manuell statt. Die Relevanz eines Dokuments wird grundsätzlich manuell durch einen Fachvertreter bestimmt. Beim Marktdateninformationssystem werden die erfassten Internetdokumente zunächst automatisiert, um die Formatierungen bereinigt und durch das entwickelte Multilayer Perceptron bewertet. Dabei werden aber nicht nur aus dem Trainingsdatenbestand generierte Deskriptoren, sondern zusätzlich ergänzende Deskriptoren der Fachabteilungsmitarbeiter verglichen. Darüber hinaus erfolgt eine Qualitätsbewertung der Internetdokumente. Die Einzelergebnisse werden zusammengefasst und bei Erreichung des unteren Schwellenwertes erfolgt die Speicherung des Dokuments im Data Warehouse. Im Gegensatz zu den anderen Konzepten verbleiben zunächst die ungeeigneten Dokumente in der Data Barn, um als Trainingsdatenbestand für das MLP genutzt zu werden. Dies erfolgt bei den anderen Konzepten nicht.

Hackathorn weist für die Datenspeicherung explizit darauf hin, dass eine objektrelationale Datenbank zu Grunde zu legen ist, in der ein Star Schema implementiert wird. Eine solche objektrelationale Datenbank wird auch für das Marktdateninformationssystem verwendet, da Internetdokumente die objektorientierte Speichermöglichkeit bedingen. Hierzu sind für den Redaktionsleitstand keine Informationen zu erhalten.

Die Daten werden den Anwendern in den Konzepten auf unterschiedliche Art und Weise zur Verfügung gestellt. Das Web Farming unterscheidet grundsätzlich zwei Distributionsarten. Eine Möglichkeit ist, die Dokumente in Form eines Push-Dienstes dem jeweiligen Personenkreis verfügbar zu machen. Dies kann durch einen E-Mail-Informationsdienst, einen Newsletter oder durch manuelle Verteilung des Content Broker geschehen. Die andere Möglichkeit ist ein Pull-Dienst, bei dem die interessierten Mitarbeiter die Dokumente auf einer Intranetseite aufrufen. Der Redaktionsleitstand besitzt eine Oberfläche, die in ihrer Funktionalität an eine Internetseite erinnert. Es handelt sich dabei um eine zentrale Seite, von der aus unterschiedliche Themenbereiche aufgerufen werden können. Die einzelnen Dokumente werden als Themenliste zu den Berichtskomplexen an-

geboten. Beim Marktdateninformationssystem ist nicht nur der jeweilige Berichtskomplex als Grafik dargestellt, sondern auch die dazu passenden Internetdokumente. Beide Darstellungen orientieren sich an ein- und derselben *Zeitlinie*. Aus der Grafik der externen Daten ist für den Anwender die Güte der Dokumente zu erkennen. Im Gegensatz zu den beiden anderen Systemen wird auf manuelle Eingaben vollständig verzichtet. So werden fehlerhafte Eingaben der Anwender ausgeschlossen und das System kann intuitiv über die Maussteuerung beeinflusst werden. Tabelle 1 fasst die zuvor beschriebenen Ergebnisse zur Integration externer Daten für das Web Farming (WebF), den Redaktionsleitstand (RLS) und das Marktdateninformationssystem (MAIS) zusammen [Feld02, S. 163].

	WebF	RLS	MAIS
Datensuche	Manuelle Suche (Surfen) im Internet Parameterlistenübertragung an Suchmaschine	Manuelle Suche (Surfen) im Internet (Pull-Dienst) Aktive Info.-Lieferanten (Push-Dienst)	Meta Suchmaschine mit automatischer Parameterübergabe aus Data Dictionary
Datenübertragung	Copy & Paste Softwareagenten für gleichbleibende Informationsquellen	Copy & Paste	HTTP-Aufruf von Servern und Seitenübertragung in Data Barn
Datenbewertung	Manuelle Prüfung	Bewertung anhand einer Deskriptorenliste Zuordnung an interne Berichte (i. A. manuell) Manuelle Bewertung	MLP
Datenspeicherung	Objektrelationale Datenbank Star Schema	unbekannt	Objektrelationale Datenbank Star Schema

Datenpräsentation	Intranetseite	Aufbau der Oberfläche orientiert sich an Internetfunktionalitäten (zentrale Seite mit Verzweigungen)	Visualisierungskonzept mit Starfield Display
	E-Mail-Dienst		
	Newsletter		
	Manuelle Verteilung durch Content Broker		

Tabelle 1: Konzeptpositionierung

Kennzeichnend für das Marktdateninformationssystem ist, dass auf verfügbare Technologien zurückgegriffen wurde und trotzdem eine völlige Automatisierung erreicht wurde, die so bei den anderen Konzepten nicht gegeben ist.

5 Fazit

Durch die Entwicklung von Marktinformationssystemen ist die Zusammenführung interner und externer Daten gelungen. Jedoch müssen die Energiedienstleistungsunternehmen zunächst einmal flächendeckend Analytische Informationssysteme in ihren Unternehmen einsetzen. Nur so lassen sich Erfahrungen sammeln, was letztendlich ein Analytisches Informationssystem im Energiehandel charakterisiert und wie es sich idealerweise realisieren lässt. Grundsätzlich ermöglicht eine solche Datenbasis den Entscheidungsträgern verbesserte Analysemöglichkeiten, um Tendenzen im Energiemarkt frühzeitig zu erkennen. Dabei ist jedoch zu beachten, dass das jeweilige Modell und das System eine hohe Flexibilität bei Anpassungen an veränderte Rahmenbedingungen im Energiemarkt gewähren muss. Es verbleibt für die Entwicklung solcher Systeme die Aufgabe, den Energiemarkt und die Handlungen der Analysten zu beobachten. Für die Weiterentwicklung des Informationssystems gilt es, die Bewertung externer Informationen zu optimieren sowie den vorgeschlagenen, weitestgehend automatisierten Integrationsprozess vollständig ohne menschliche Eingriffe ausführen zu lassen.

Literatur

- [Ahl⁺99] Ahlberg, C.; Shneidermann, B.: Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays, in: Card, S. K.; Mackinlay, J. D.; Shneiderman, B. (Hrsg.): Readings in Information Visualization - Using Vision to Think, San Francisco, 1999, S. 244 – 252.
- [Balz01] Balzert, H.: Lehrbuch der Software-Technik I - Software-Entwicklung, 2. Auflage, Heidelberg [u. a.], 2001.

- [Bene98] Benenati, I.: Neuronale Netze im Portfoliomanagement, Wiesbaden, 1998.
- [Bish95] Bishop, C. M.: Neural Networks for Pattern Recognition, Oxford [u. a.], 1995.
- [Brau95] Brause, R.: Neuronale Netze: eine Einführung in die Neuroinformatik, 2. Überarbeitete und erweiterte Auflage, Stuttgart, 1995.
- [Bulo98] Bulos, D.: OLAP Database Design - A New Dimension, in: Chamoni, P.; Gluchowski, P.: Analytische Informationssysteme - Einordnung und Überblick, in: Chamoni, P.; Gluchowski, P. (Hrsg.): Analytische Informationssysteme - Data Warehouse, On-Line Analytical Processing, Data Mining, Berlin, 1998, S. 251 - 261.
- [Car⁺99] Card, S. K.; Mackinlay, J. D.; Shneiderman, B. (Hrsg.): Readings in Information Visualization - Using Vision to Think, San Francisco, 1999.
- [Chen76] Chen, P. P. S.: The Entity Relationship Model - Toward a Unified View of Data, in: ACM Transactions on Database Systems, Vol. 1, No. 1, March, 1976, S. 9 - 36.
- [Elm⁺94] Elmasri, R. A.; Navathe, S. B.: Fundamentals of Database Systems, 2nd ed., Redwood City [u. a.], 1994.
- [Feld02] Felden, C.: Konzept zum Aufbau eines Marktdateninformationssystems für den Energiehandel auf der Basis interner und externer Daten, Wiesbaden, 2002.
- [Gehr00] Gehrke, C.: Informationsagenten im Data Warehousing, Heidelberg, 2000.
- [Glu⁺97] Gluchowski, P.; Gabriel R.; Chamoni, P.: Managament Support Systeme - Computerunterstützte Informationssysteme für Führungskräfte und Entscheidungsträger, Berlin [u. a.], 1997.
- [Hack98] Hackathorn, R. D.: Web Farming for the Data Warehouse, San Francisco, 1998.
- [Hahn99] Hahne, M.: Logische Datenmodellierung für das Data Warehouse - Bestandteil und Varianten des Star Schemas, in: Chamoni, P.; Gluchowski, P.: Analytische Informationssysteme - Data Warehouse, On-Line Analytical Processing, Data Mining, 2. Auflage; Berlin [u. a.], 1999, S. 145 - 170.
- [Inm⁺00] Inmon, W. H.; Terdeman, R. H.; Imhoff, C.: Exploration Warehousing - Turning Business Information into Business Opportunity, New York [u. a.], 2000.
- [Khan90] Khanna, T.: Foundations of Neural Networks, Reading, Massachusetts [u. a.], 1990.
- [Kinn94] Kinnebrock, W.: Neuronale Netze: Grundlagen, Anwendungen, Beispiele, 2. Auflage, München [u. a.], 1994.
- [Küng94] Küng, P.: Datenbanksysteme: Entwicklungsstand, Anforderungen und Bedeutung neuerer Konzepte, Zürich, 1994.
- [Mazz92] Mazzetti, A.: Praktische Einführung in neuronale Netze, Hannover, 1992.
- [McC⁺90] McCord Nelson, M.; Illingworth, W. T.: A practical guide to neural nets, Reading [u. a.], 1990.
- [McC⁺43] McCulloch, W. S.; Pitts, W.: A logical calculus of the ideas immanent in nervous activity, in: Bulletin of mathematical biophysics, Vol. 5, 1943, S. 115 - 133.

- [Meie00] Meier, M.: Integration externer Daten in Planungs- und Kontrollsysteme - Ein Redaktionsleitstand für Informationen aus dem Internet, Wiesbaden, 2000.
- [Naye93] Nayer, M.: Achieving Information Integrity: A Strategic Imperative, in: Information Systems Management, Vol. 10, No. 2, 1993, S. 51 - 58.
- [Niel93] Nielsen, J.: Usability Engineering; Boston [u. a.], 1993.
- [Raa⁺00] Raab, R.; Bach, V.; Österle, H.: Systematische Erschließung der Informationsquelle Internet, in: industrie Management - Life Cycle Management, 1/2000.
- [Rand82] Randell, B. (ed.): The origins of digital computers: selected papers, 3rd printing, Berlin [u. a.], 1982.
- [seve03] o. V.: Five Letters – Was versteckt sich hinter den fünf Buchstaben, URL: http://www.seven2one.de/five_letters.htm, Abruf: 2003-02-14.
- [Stah95] Stahlknecht, P.: Einführung in die Wirtschaftsinformatik, 7. Auflage, Berlin [u. a.], 1995.
- [Stoc01] Stock, S.: Modellierung zeitbezogener Daten im Data Warehouse, Wiesbaden, 2001.
- [Stre96] Streubel, F.: Theoretische Fundierung eines ganzheitlichen Informationsmanagements, Arbeitsbericht des Lehrstuhls für Wirtschaftsinformatik 96-21, Ruhr-Universität Bochum, Bochum, 1996.
- [Teut01] Teuteberg, F.: Agentenbasierte Informationserschließung im World Wide Web unter Einsatz von Künstlichen Neuronalen Netzen und Fuzzy-Logik, Lohmar, Köln, 2001.
- [Uhr⁺98] Uhr, W.; Breuer, S.-E. (Hrsg.): Integration externer Informationen in Management Support Systems, Wirtschaftsinformatik Fachtagung, Technische Universität Dresden, Oktober 1998, S. 8 - 9.
- [vonN45] von Neumann, J.: First draft of a report on the EDVAC, Moore School of Electrical Engineering, University of Pennsylvania, 1945.
- [Wert86] Wertz, C. J.: The Data Dictionary - Concepts and Uses, Amsterdam, 1986.
- [Woo⁺94] Wooldridge, M. J.; Jennings, N. R.: Intelligent Agents - ECAI-94 Workshop on Agent Theories, Architectures, and Languages, Amsterdam, The Netherlands, August 8 - 9, 1994, Proceedings, Berlin [u. a.], 1994.
- [Zarn99] Zarnekow, R.: Softwareagenten und elektronische Kaufprozesse - Referenzmodelle zur Integration, Wiesbaden, 1999.
- [Zaun99] Zaun, D. P.: Künstliche neuronale Netze und Computerlinguistik, Tübingen, Köln, 1999.